STATISTICS OF CLOSEST RETURN FOR SOME NON UNIFORMLY HYPERBOLIC SYSTEMS.

P.COLLET

Centre de Physique Théorique
Laboratoire CNRS UMR 7644
Ecole Polytechnique
F-91128 Palaiseau Cedex (France)

Abstract: For non uniformly hyperbolic maps of the interval with exponential decay of correlations we prove that the law of closest return to a given point when suitably normalized is almost surely asymptotically exponential. A similar result holds when the reference point is the initial point of the trajectory. We use the framework for non uniformly hyperbolic dynamical systems developed by L.S.Young.

Keywords: entrance time, extreme statistics, decay of correlations.

MSC: 34C35, 60G70

I. INTRODUCTION.

The statistics of entrance time in a small set is a long standing important problem. One of the first instances is the famous question raised by Boltzmann about the time it takes to observe the clustering of all the molecules of a gas in only half of the available volume. Other important applications include the occurrence of rare events, eventually catastrophic. More recently, several algorithms were proposed to measure various quantities like dimension and entropy of dynamical systems which involve such entrance time questions. We refer to [ABST] for a review of these algorithms. Another application concerns the optimal compression of data sets. The well known compression algorithm developed by Ziv and Wyner is based on the coding of repetitions of patterns previously coded. Its optimality is based on almost sure results for the typical recurrence time which were proven for general ergodic sources by Ornstein and Weiss. We refer to [WZW] and [Sh] for recent reviews on this subject and references to older works. This algorithm can also be viewed as a way of measuring the entropy.

In all these questions, the asymptotic result is determined by a law of large numbers. One would like to understand the fluctuations in order to control the rate of convergence and for statistical purposes. The rate of convergence for the Wyner-Ziv algorithm was recently derived for the case of sufficiently mixing sources in [CGS], [K] and [P].

In the present paper we derive similar results from a different point of view which emphasizes the topology of the phase space, and for systems which are non uniformly hyperbolic. The problem discussed below can be formulated in terms of extreme statistics (see [G]). Consider a dynamical system given by a (compact) phase space Ω equipped with a metric d, a continuous map T on Ω and a T invariant ergodic probability measure μ. Assume a point x in phase space has been chosen and define a sequence (X_j) of real random variables on Ω by

$$X_j(y) = -\log(d(x, T^j(y))) \ .$$

Let (Z_n) be the sequence of successive maxima of the sequence (X_j), namely

$$Z_n(y) = \sup_{0 \le j \le n} X_j(y) \ .$$

One may conjecture that if the dynamical system satisfies the Eckmann-Ruelle conjecture, then almost surely

$$\lim_{n \to \infty} \frac{Z_n}{\log n} = D^{-1}$$

where D is the dimension of the measure. A more precise question is to to ask if there are two sequences of positive numbers (a_n) and (b_n) such that for any positive real number s, the sequence

$$\mathbb{P}\left(Z_n > a_n s + b_n\right)$$

converges. We will denote by $\mathbb{P}$ the probability associated to the measure μ and by $\mathbb{E}$ the corresponding expectation. We refer to [G] for the study of the statistics of extremes for independent random variables and for related questions. In the present case of dynamical systems, one may expect that the result (and in particular the choice of the two sequences (a_n) and (b_n)) depends on the point x chosen at the beginning. We will assume below

that the point x has been chosen at random with respect to μ and we will prove results almost surely with respect to this choice. In order to simplify the notation we will usually not mention the x dependence since this is a point which is chosen once for all. When this dependence needs to be emphasized we will denote it by an exponent in (X_j^x) and (Z_n^x).

We now describe the dynamical systems for which the result will be proven. An abstract frame for non uniformly hyperbolic systems was introduced by L.S.Young in [Y1] and [Y2] (see also references therein and [BV] and [KN] for earlier constructions). Instead of using the completely abstract formulation we will rather keep the equivalent version in the phase space. The reason for doing so is that the topology is somewhat obscured when the system is lifted to the abstract context. We will also work explicitly the case of maps of the interval with exponential decay of correlations although several results extend to more general situations. We will mention some of these extensions when appropriate. We now formulate the hypothesis on our dynamical system which follow directly from the work of L.S.Young.

We will consider a C^2 map f of the interval $[a,b]$ into itself and we denote by K the sup norm of its derivative

$$K = \sup_{x\in[a,b]} |f'(x)| \ .$$

We assume there is an open interval Λ in $[a,b]$ with dense orbit and with the following properties.

H1 *There exists a sequence $(R_i)_{i\in\mathbf{N}}$ of positive integers, with largest common divisor equal to one and a sequence of disjoint open subintervals $(\Lambda_i)_{i\in\mathbf{N}}$ of Λ satisfying $\lambda(\Lambda\backslash\cup_i\overline{\Lambda}_i)=0$ such that the following holds. For any $j\in\mathbf{N}$, f^{R_j} is a bijection from Λ_j to Λ. There exits also an integer valued function s defined on $\Lambda\times\Lambda$ such that for x and y in Λ, the orbits of x and y follow each other up to time $s(x,y)$ in the sense that the corresponding orbits under f^R fall in the same Λ_i. We assume s is finite $\lambda\times\lambda$ almost surely, where λ is the Lebesgue measure.*

Hypothesis **H1** is of course of Markov type. Note however that the number R_j may not be the first return to Λ of Λ_j. It is a return where the properties of uniform backward contraction **H2** and uniform distortions **H3** are satisfied. We will speak of these returns as "official" returns.

H2 *There is a constant $C>0$ and a number $0<\beta<1$ such that for any x and y in Λ, and any $0\le n\le s(x,y)$ we have*

$$|f^n(x)-f^n(y)|\le C\beta^{s(x,y)-n} \ .$$

H3 *For any x, y in Λ and any $0\le k\le n\le s(x,y)$ we have*

$$\log\left(\prod_{i=k}^{n}\frac{|f'(f^i(x))|}{|f'(f^i(y))|}\right)\le C\beta^{s(x,y)-n} \ .$$

In [Y1] and [Y2] examples of dynamical systems where given where these hypothesis are satisfied. In particular unimodal maps of the interval with sufficient instability of the

critical orbit satisfy these hypothesis. There are also examples with neutral fixed points and higher dimensional cases. We will make some comments below about these cases.

After having described the setting in phase space, we now come to the invariant measure. Let λ denote the Lebesgue measure on $[a, b]$. The next hypothesis concerns the random variable R defined on Λ by $R(x) = R_j$ if $x \in \Lambda_j$.

H4 *The random variable R is integrable with respect to λ (restricted to Λ).*

One of the first results of L.S.Young is that under the above hypotheses, there is a measure μ_0 on Λ which is equivalent to λ (more precisely with a density bounded above and bounded below away from zero) and which is invariant and ergodic for f^R. This leads to an invariant ergodic measure μ for the map f which is given by

$$\mu(A) = Z^{-1} \sum_{l} \sum_{j=0}^{R_l - 1} \mu_0 \left(\Lambda_l \cap f^{-j}\left(A \cap f^j(\Lambda_l)\right)\right) \ , \qquad (I.1)$$

with

$$Z = \sum_l R_l \mu_0 \left(\Lambda_l\right) = \int R \, d\mu_0 \ .$$

One of the main result in [Y1] and [Y2] is a bound on the decay of correlations for for Hölder continuous. Namely if g_1 is Hölder continuous and $g_2 \in L^\infty(d\lambda)$, the decay of correlations $\alpha(\,\cdot\,)$ defined by

$$\alpha(n) = \left| \int g_1 \ g_2 \circ f^n d\mu \ - \int g_1 \ d\mu \int g_2 \ d\mu \right|$$

is related to the behaviour for large n of $\lambda(R > n)$. If this sequence decays exponentially fast, the same is true for α with the same decay rate. A similar result holds in the case of polynomial decay. In the case of exponential decay of $\lambda(R > n)$, a stronger version of the decay of correlations was proven in [Y1] which is analogous to the case of piecewise expanding maps of the interval. Although this stronger result would slightly simplify some arguments below, we will not use it. We can now formulate our main result.

Theorem I.1. *Assume the hypotheses* **H1**-**H4**. *Assume $\lambda(R > k)$ decays exponentially fast with k. Then for μ (or Lebesgue) almost every x we have*

$$\lim_{n\to\infty} \mathbb{P}\left(Z_n^x < s + \log n\right) = e^{-2h(x)e^{-s}}$$

where h is the density of the absolutely continuous invariant measure μ.

This is sometimes called Gumbel's law. There is an obvious relation with the entrance time in a ball of radius e^{-s}/n centered at x, i.e. if $Z_n < s + \log n$, the orbit has not entered the ball up to time n.

In the next section we will prove some preparatory results, and in particular we will control the measure of the set of points which recur too fast. The proof of Theorem I.1 will then be given in section 3, inspired by the techniques used for extreme statistics. In

section 4, we will establish a fluctuation result for the case where the initial point is the reference point.

We mention also that some intermediate results proven below where already derived in explicit situations in order to prove hypotheses **H1**-**H4** or the decay of $\lambda(R>k)$. One of the goal of this paper is to show that the previously mentioned hypotheses are sufficient to prove the result without reference to particular constructions.

II. ESTIMATES FOR THE SET OF RAPIDLY RECURRING POINTS.

In this section we will estimate the measure of some sets of points with exceptional behaviour. For later references we start with the following easy lemma.

Lemma II.1.

$$\sum_{l\,,\,R_l>k} R_l\;\lambda(\Lambda_l)\le\begin{cases}2\sum_{s=k/2}^{\infty}\;\lambda(\{R>s\})\;,\\ \sum_{s=k}^{\infty}\;\lambda(\{R>s\})+k\,\lambda(\{R>k\})\;.\end{cases}$$

Proof. We have indeed for any $q>0$

$$\sum_{s=q}^{\infty}\lambda(\{R>s\})=\sum_{s=q}^{\infty}\sum_{l=1}^{\infty}\lambda(\{R>s\}\cap\Lambda_l)$$

$$=\sum_{l,\,R_l>q}^{\infty}(R_l-q)\lambda(\Lambda_l)\;.$$

For the first estimate, we take $q=k/2$ and restrict the last sum in the above equality to the range $R_l>k$ which implies $R_l-q\ge R_l/2$ and the result follows. For the second estimate we simply take $q=k$ and rearrange the equality.

We will need later an estimate of the μ measure of sets with small Lebesgue measure.

Lemma II.2. *Assume exponential decay in k of $\lambda(\{R>k\})$. Then there are two positive constants C and θ such that for any Lebesgue measurable set I, we have*

$$\mu(I)\le C\lambda(I)^{\theta}\;.$$

Proof. From formula (I.1), for $0\le j<R_l$ we have to consider the sets

$$I_{l,j}=I\cap f^{j}(\Lambda_l)\;.$$

Since f^j is injective on Λ_l, there is a set $\tilde I_{j,l}$ in Λ_l such that $f^j(\tilde I_{j,l})=I\cap f^j(\Lambda_l)$. There are now two cases.

In the first case, we assume $K^{R_l-j}\le|I|^{-1/2}$. We now use the distortion bound on Λ_l for f^{R_l}. We get

$$\frac{|\tilde I_{j,l}|}{|\Lambda_l|}\le\mathcal{O}(1)\frac{|f^{R_l}(\tilde I_{j,l})|}{|f^{R_l}(\Lambda_l)|}=\mathcal{O}(1)\frac{\left|f^{R_l-j}\big(I\cap f^j(\Lambda_l)\big)\right|}{|f^{R_l}(\Lambda_l)|}\le\mathcal{O}(1)K^{R_l-j}|I|\le\mathcal{O}(1)|I|^{1/2}.$$

This implies

$$|\tilde{I}_{j,l}| \le \mathcal{O}(1)|I|^{1/2}|\Lambda_l| \ ,$$

and since μ_0 is equivalent to the Lebesgue measure

$$\mu_0(\tilde{I}_{j,l}) \le \mathcal{O}(1)|I|^{1/2}\mu_0(\Lambda_l) \ .$$

We can now sum over j and l to get

$$\sum_{j,\,l,\,K^{R_l-j}<|I|^{-1/2}} \mu_0(\tilde{I}_{j,l}) \le \mathcal{O}(1)|I|^{1/2}\sum_l R_l\ \mu_0(\Lambda_l) \ .$$

This last sum is finite since R is integrable with respect to λ by hypothesis **H4**, and μ_0 is equivalent to λ.

We now deal with the second case, namely $K^{R_l-j} > |I|^{-1/2}$. This implies of course

$$R_l \ge \frac{\log|I|^{-1}}{2\log K} \ .$$

Therefore

$$\sum_{j,\,l,\,K^{R_l-j}\ge|I|^{-1/2}} \mu_0(\tilde{I}_{j,l}) \ \le\ \mathcal{O}(1) \sum_{l\,,\,R_l>\frac{\log|I|^{-1}}{2\log K}} R_l\ \mu_0(\Lambda_l) \le \mathcal{O}(1) \sum_{l\,,\,R_l>\frac{\log|I|^{-1}}{2\log K}} R_l\ \lambda(\Lambda_l) \ ,$$

and the result follows from the assumption on the exponential decay of $\lambda(R>k)$ and Lemma II.1.

Remark. Lemma II.2 implies that the density h of the measure μ with respect to the Lebesgue measure belongs to some L^p with $p>1$. This can also be proven directly using estimates similar to those in the above proof. Some examples of maps of the interval with neutral fixed point are known to have an invariant measure with a density in some L^p (see [T]) while the bound on $\lambda(R>k)$ is only known to be polynomial and the above proof does not work in that case.

The proof of Theorem I.1 in the next section will require that the point x is not too rapidly recurrent. We will now prove that rapidly recurrent points are exceptional with respect to the measure μ. It is convenient to define for any integer k, and any positive number ϵ the set $\mathcal{E}_k(\epsilon)$ by

$$\mathcal{E}_k(\epsilon) = \{x\,,\ |x-f^k(x)|<\epsilon\} \ .$$

Proposition II.3. *There exists positive constants C, α and $\eta<1$ such that for any integer k and any $\epsilon>0$ we have*

$$\mu(\mathcal{E}_k(\epsilon)) \le C\left(k^2\epsilon^\eta + e^{-\alpha k}\right) \ .$$

Proof. We will consider the intersection with $\mathcal{E}_k(\epsilon)$ of the various intervals of monotonicity of f^k. From formula (I.1), we have to consider the intersection of these sets with $f^j(\Lambda_l)$.

We will start by choosing a number $\zeta > 0$ such that $\beta K^{2\zeta} < 1$ and assume first that $R_l < \zeta k$. If we apply $f^{R_l - j}$ on $f^j(\Lambda_l)$, we land in Λ and we have to apply f^{k-R_l+j}. At this point it is convenient to introduce the following construction. Let (s_j) be a sequence of integers. We denote by $\Lambda_{s_1,\, s_2,\, \cdots,\, s_r}$ the set

$$\Lambda_{s_1,\, s_2,\, \cdots,\, s_r} = \Lambda_{s_1} \cap f^{-R_{s_1}} \Lambda_{s_2} \cap f^{-(R_{s_1}+R_{s_2})} \Lambda_{s_3} \cap \ \cdots\ \cap f^{-(R_{s_1}+\,\cdots\,+R_{s_{r-1}})} \Lambda_{s_r} \ .$$

In other words, this is the subset A of Λ_{s_1} which is mapped by $f^{R_{s_1}+\cdots+R_{s_{r-1}}}$ bijectively on Λ_{s_r} with

$$f^{R_{s_1}+\,\cdots\,+R_{s_p}}(A) \subset \Lambda_{s_{p+1}} \ .$$

for $p = 1, \ \cdots\ , \ r-1$.

For fixed k, l and j, we now consider all the sets $\Lambda_{s_1,\,\cdots\,,\, s_r}$ with $R_{s_1} + \ \cdots\ + R_{s_{r-1}} + R_l - j < k$ and $R_{s_1} + \cdots + R_{s_r} + R_l - j \geq k$. Together with $\{R > k-1-R_l+j\}$, this gives a partition of Λ. We then construct a partition of $f^j(\Lambda_l)$ by pulling back this partition by f^{R_l-j}. We now consider f^k on each atom of this partition. Let

$$I = I_{l,j,s_1,\cdots,s_r} = f^j(\Lambda_l) \cap f^{j-R_l}\left(\Lambda_{s_1,\,\cdots\,,\, s_r}\right) \ .$$

By construction, f^k is injective on the set I and we have controled distorsion by **H3**. We now prove that the slope of f^k is uniformly larger than two for k large enough and $R_{s_r} < \zeta k$. Let $\tilde{I}$ be the segment contained in Λ_l which is mapped bijectively by f^j on I. By contraction and distorsion, we have

$$|\Lambda| = |f^{R_l+R_{s_1}+\,\cdots\,+R_{s_r}}(\tilde{I})| = \mathcal{O}(1)\ \left|\left(f^{R_l+R_{s_1}+\,\cdots\,+R_{s_r}}\right)'\Big|_{\tilde{I}}\right|\ |\tilde{I}|$$

$$\leq \mathcal{O}(1)\ \left|\left(f^{R_l+R_{s_1}+\,\cdots\,+R_{s_r}}\right)'\Big|_{\tilde{I}}\right|\ \beta^{R_l+R_{s_1}+\,\cdots\,+R_{s_r}}$$

$$\leq \mathcal{O}(1)\ \left|f^{k'}\Big|_I\right|\ \beta^k\ K^j\ K^{R_l-j+R_{s_1}+\,\cdots\,+R_{s_r}-k}$$

$$\leq \mathcal{O}(1)\ \left|f^{k'}\Big|_I\right|\ \beta^k\ K^{2\zeta k}$$

if we assume $R_{s_r} < \zeta k$. The result now follows for k large enough since $\beta K^{2\zeta} < 1$.

From this result it follows easily that if $\mathcal{E}_k(\epsilon) \cap I$ is not empty, then it is a segment denoted below by J. Assume first that I has a "large" image under f^k, namely

$$|f^k(I)| \geq \delta \ ,$$

where δ is a positive number to be chosen adequately later on.

Since J is a segment, it follows easily from the definition of $\mathcal{E}_k(\epsilon)$ that $|f^k(J)| \leq 4\epsilon$ if we assume $|f^{k'}| > 2$. Using distorsion, we get

$$|J|/|I| \leq \mathcal{O}(1)\epsilon/\delta \ .$$

Using again distorsion, we get

$$|\Lambda_l \cap f^{-j}(J)|/|\Lambda_l \cap f^{-j}(I)| \le \mathcal{O}(1)\epsilon/\delta \; .$$

This implies since μ_0 is equivalent to the Lebesgue measure on Λ

$$\mu_0(\Lambda_l \cap f^{-j}(J)) \le \mathcal{O}(1)\frac{\epsilon}{\delta}\mu_0(\Lambda_l \cap f^{-j}(I)) \; .$$

We can now sum over all the above intervals I contained in $f^j(\Lambda_l)$ and with "large" image. Since they are disjoint we get a contribution bounded above by $\mathcal{O}(1)(\epsilon/\delta)\mu_0(\Lambda_l)$. Summing over j we get a bound $\mathcal{O}(1)(\epsilon/\delta)R_l\mu_0(\Lambda_l)$. Summing over l we get finally an estimate $\mathcal{O}(1)(\epsilon/\delta)$. This ends the estimate in the good case of segments I with "large" images. We now have to collect the estimates for all the left-over bad cases.

First of all we have assumed $R_l \le \zeta k$. We have

$$\sum_{l\,,\,R_l>\zeta k}\;\sum_{j=0}^{R_l-1}\mu_0\left(\Lambda_l \cap f^{-j}\left(\mathcal{E}_k(\epsilon)\cap f^j(\Lambda_l)\right)\right) \le \sum_{l\,,\,R_l>\zeta k} R_l\mu_0(\Lambda_l)$$

and we have a bound from Lemma II.1.

We now deal with the bad cases associated to R_{s_r}. We have by definition

$$\mathcal{E}_k(\epsilon)\cap f^j(\Lambda_l) =$$

$$\bigcup_{R_{s_1}+\cdots+R_{s_{r-1}}<k\le R_{s_1}+\cdots+R_{s_r}} \left(I_{j,l,s_1,\cdots,s_r}\cap\mathcal{E}_k(\epsilon)\right)$$

$$\bigcup\left(\mathcal{E}_k(\epsilon)\cap f^j(\Lambda_l)\cap f^{-R_l+j}(\{R>k-R_l+j\})\right) \; .$$

We first consider the last set. We have to estimate the μ_0 measure of

$$\Lambda_l\cap f^{-j}\left(\mathcal{E}_k(\epsilon)\cap f^j(\Lambda_l)\cap f^{-R_l+j}(\{R>k-R_l+j\})\right) \; .$$

This set is obviously contained in

$$\Lambda_l\cap f^{-j}\left(f^{-R_l+j}(\{R>k-R_l+j\})\cap f^j(\Lambda_l)\right) \; ,$$

which is a subset of

$$\Lambda_l\cap f^{-j}\left(f^{-R_l+j}(\{R>(1-\zeta)k\})\right) = \Lambda_l\cap f^{-R_l}(\{R>(1-\zeta)k\}) \; ,$$

if $R_l\le\zeta k$ (recall that $R_l>j$). We get

$$\sum_{l,\,R_l<\zeta\,k}\;\sum_{0\le j<R_l}\mu_0\left(\Lambda_l\cap f^{-j}\left(\mathcal{E}_k(\epsilon)\cap f^j(\Lambda_l)\cap f^{-R_l+j}(\{R>k-R_l+j\})\right)\right)$$

$$\leq \sum_{l,\, R_l<\zeta\, k} \quad \sum_{0\leq j<R_l} \mu_0\left(\Lambda_l \cap f^{-R_l}(\{R>(1-\zeta)k\})\right)$$

$$\leq k \sum_{l,\, R_l<\zeta\, k} \mu_0\left(\Lambda_l \cap f^{-R_l}(\{R>(1-\zeta)k\})\right) \ .$$

By distorsion, we have

$$\frac{\left|\Lambda_l \cap f^{-R_l}(\{R>(1-\zeta)k\})\right|}{|\Lambda_l|} \leq$$

$$\mathcal{O}(1)\left|f^{R_l}\left(\Lambda_l \cap f^{-R_l}(\{R>(1-\zeta)k\})\right)\right| \leq \mathcal{O}(1)|\{R>(1-\zeta)k\}| \ .$$

Therefore

$$\sum_{l,\, R_l<\zeta\, k}\sum_{0\leq j<R_l} \mu_0\left(\Lambda_l \cap f^{-j}\left(\mathcal{E}_k(\epsilon)\cap f^{-R_l+j}(\{R>k-R_l+j\})\right)\right) \leq k\lambda(R>(1-\zeta)k) \ .$$

We now consider the case $R_{s_r}>q$ for some integer q. We have

$$\bigcup_{R_{s_1}+\cdots+R_{s_{r-1}}<k\leq R_{s_1}+\cdots+R_{s_r}\ ,\ R_{s_r}>q} \left(I_{j,l,s_1,\cdots,s_r}\cap\mathcal{E}_k(\epsilon)\right)$$

$$\subset f^j(\Lambda_l)\bigcap\left(\bigcup_{m=0}^{k} f^{-m}(\{R>q\})\right) \ .$$

We recall that $\{R>q\}$ is a subset of Λ. Applying f^{-j} and intersecting with Λ_l, we have to estimate

$$\sum_{l\,,\, R_l<\zeta k}\ \sum_{j=0}^{R_l-1}\sum_{m=0}^{k}\mu_0\left(\Lambda_l\cap f^{-j-m}(\{R>q\})\right) \ .$$

We can now use the fact that on Λ we have $\mu\geq\mu_0$. Therefore, the above quantity is bounded by

$$\sum_{l\,,\, R_l<\zeta k}\ \sum_{j=0}^{R_l-1}\sum_{m=0}^{k}\mu\left(\Lambda_l\cap f^{-j-m}(\{R>q\})\right) \leq \sum_{l\,,\, R_l<\zeta k}\ \sum_{j=0}^{R_l-1}\sum_{m=0}^{2k}\mu\left(\Lambda_l\cap f^{-m}(\{R>q\})\right)$$

$$\leq k\sum_{l\,,\, R_l<\zeta k}\sum_{m=0}^{2k}\mu\left(\Lambda_l\cap f^{-m}(\{R>q\})\right) \leq k\sum_{m=0}^{2k}\mu\left(f^{-m}(\{R>q\})\right)\leq 3k^2\mu(\{R>q\}) \ .$$

In particular, we have

$$\mu\left(\left\{R_{s_r}>\zeta k\right\}\right)\leq 3k^2\mu(\{R>\zeta k\}) \ .$$

We now have to deal with the cases $|f^k(I)|<\delta$. Let

$$p = R_l - j + R_{s_1} + \cdots + R_{s_r} - k \ .$$

In other words, p is the number of iterations needed from $f^k(I)$ to reach $f^{R_{s_r}}(\Lambda_{s_r}) = \Lambda$, hence

$$\Lambda = f^p(f^k(I)) \ .$$

Therefore

$$|\Lambda| \le K^p |f^k(I)| \le K^p \delta \ ,$$

which implies

$$p \ge \mathcal{O}(1) \log \delta^{-1}$$

and therefore

$$R_{s_r} \ge p \ge \mathcal{O}(1) \log \delta^{-1} \ .$$

We now collect all the estimates and get

$$\mu(\mathcal{E}_k(\epsilon)) \le$$

$$\mathcal{O}(1) \left(\frac{\epsilon}{\delta} + \sum_{s > \zeta k/2} \mu_0(R > s) + k\mu_0(R > \zeta k) + k^2 \mu(R > \zeta k) + k^2 \mu\big(R > \mathcal{O}(1) \log \delta^{-1}\big) \right)$$

This can be expressed in terms of λ only using Lemma II.2.

If we assume that $\lambda(R > k)$ decays exponentially fast, namely

$$\lambda(R > k) \le \mathcal{O}(1) e^{-\alpha' k}$$

for some $\alpha' > 0$, we get

$$\mu(\mathcal{E}_k(\epsilon)) \le \mathcal{O}(1) \left(\frac{\epsilon}{\delta} + k e^{-\alpha' \zeta k} + k^2 \delta^{\gamma} \right)$$

for some $1 > \gamma > 0$. The result follows by taking the minimum with respect to δ.

In the above proof, one can avoid using the invariant measure μ in the estimate, using instead the measure μ_0 invariant by the map f^R. This allows to use the same method in higher dimensional situations. The good case corresponds to "large" enough local unstable manifolds and give a relative bound of order ϵ/δ which can be integrated against the transverse measure. The bad cases are then handled by showing that they all correspond to large values of R.

We now derive several consequences of Proposition II.3. Let (E_k) be the sequence of sets defined by

$$E_k = \left\{ y \,|\, \exists\, j \ 1 \le j \le (\log k)^5 \ , \ |y - f^j(y)| \le k^{-1} \right\} \ .$$

Corollary II.4. *There exists positive constants C' and $\beta' < 1$ such that for any integer k*

$$\mu(E_k) \le C' \, k^{-\beta'} \ .$$

Proof. Note first that the estimate in Proposition II.3 is not very good for small k. This can be improved as follows. We observe that since f has a slope bounded in absolute value by K, the inequality

$$|f^j(x) - x| \le \epsilon$$

implies

$$|f^{2j}(x) - x| \le |f^{2j}(x) - f^j(x)| + |f^j(x) - x| \le (K^j + 1)\epsilon \ ,$$

and more generally for any $r \ge 1$

$$|f^{rj}(x) - x| \le (K^j + 1)^{r-1}\epsilon \ .$$

In other words

$$\mathcal{E}_j(\epsilon) \subset \mathcal{E}_{rj}((K^j + 1)^{r-1}\epsilon) \ .$$

This implies together with Proposition II.3 that for any $r \ge 1$ we have

$$\mu(\mathcal{E}_j(\epsilon)) \le \mathcal{O}(1)\left((K^j + 1)^{(r-1)\eta}(rj)^2\epsilon^\eta + e^{-\alpha r j}\right) \ .$$

Taking the minimum with respect to r, it follows that there are two constants $C'' > 0$ and $\beta'' > 0$ such that for any integer j

$$\mu(\mathcal{E}_j(\epsilon)) \le C''\epsilon^{\beta''} j^2 \ .$$

The result follows by choosing $\epsilon = 1/k$ and summing over j from 1 to $(\log k)^5$.

Remark. By a similar argument and using the Borel-Cantelli Lemma, one can show that there is number $\rho > 0$ such that the set of x for which the event $|x - f^k(x)| \le k^{-\rho}$ occurs for infinitely many k is of measure zero. This would be enough for the proof in section 3 if we use the stronger form of the decay of correlations mentioned in the introduction.

In order to be able to use only the weaker form of the decay of correlations, we are going to straighten the above estimate. We will not only control the set of points which recur too fast but also the set of points for which a neighbor recur too fast.

For positive number ψ and ρ to be fixed below, we define a sequence of measurable sets (F_k) by

$$F_k = \left\{x \,|\, \mu\big([x - k^{-\psi}, x + k^{-\psi}] \cap E_{k^\psi}\big) \ge 2\,k^{-(1+\rho)\psi}\right\} \ .$$

Lemma II.5. *The exists positive numbers ρ and ψ such that the set of x which belong to infinitely many F_k is of Lebesgue measure zero (and consequently of μ measure zero).*

Proof. We will first prove that for a suitable choice of ρ and ψ the sequence $(\lambda(F_k))$ is summable.

Let χ_{E_n} denote the characteristic function of E_n. We have already observed that as a consequence of Lemma II.2, the density h of μ belongs to $L^p([a, b], d\lambda)$ for some $p > 1$. Therefore the function $h\chi_{E_n}$ belongs also to this space. Moreover using Hölders inequality

and Corollary II.4 its $L^{p'}$ norm with $p' = (1+p)/2$ is bounded above by $\mathcal{O}(1)n^{-\vartheta}$ for some $\vartheta > 0$. We now introduce the maximal function M_n defined by

$$M_n(x) = \sup_{a>0} \frac{1}{2a} \int_{x-a}^{x+a} h(y)\, \chi_{E_n}(y)\, dy \;.$$

By a well known result of Hardy and Littlewood (see [St]), this function also belongs to $L^{p'}$ and has a norm bounded above by $\mathcal{O}(1)n^{-\vartheta}$. In particular it follows from the inequality of Chebyshev that

$$\lambda\left(M_n \geq n^{-\vartheta/2}\right) \leq \mathcal{O}(1)n^{-p'\vartheta/2} \;.$$

In other words if $\rho = \vartheta/2$ and $\psi > 4/(p'\vartheta)$ we have (for k large enough)

$$F_k \subset \left\{M_{k^\psi} \geq k^{-\psi\vartheta/2}\right\}$$

which implies

$$\lambda(F_k) \leq \mathcal{O}(1)k^{-\psi p'\vartheta/2} \leq \mathcal{O}(1)k^{-2} \;.$$

This last quantity is summable over k and the result follows at once from the Borel-Cantelli Lemma.

III. PROOF OF THEOREM I.1.

The strategy is inspired by the technique of extreme statistics, see for instance [G]. We briefly explain how it works. Assuming $n = pq$ with $p \approx \sqrt{n}$ and choosing $s \approx (\log n)^2$ we show that for $u_n = v + \log n$

$$\mathbb{P}(Z_n < u_n) \approx \mathbb{P}(Z_{q(p+s)} < u_n)$$

We then "dig holes" of length s separating intervals of size p. Using decay of correlations we get

$$\mathbb{P}(Z_{q(p+s)} < u_n) \approx \mathbb{P}(Z_p < u_n)^q \;,$$

and also

$$\mathbb{P}(Z_p < u_n) \approx 1 - p\mathbb{P}(X > u_n) \;.$$

As the reader can check, all the arguments in this section which do not involve the results of section II work also with a fast enough polynomial decay of correlations (with suitable choices for p and s).

It is convenient to use as much as possible set theoretic estimates as presented in the next lemma. In order to alleviate the notation, we will denote by $\{A\}$ the characteristic function of the event A.

Lemma III.1. *For any $k > 0$ we have*

$$\sum_{j=1}^{k}\{X_j > u\} \geq \{Z_k > u\} \geq \sum_{j=1}^{k}\{X_j > u\} - \sum_{j=1}^{k} \quad \sum_{l\neq j\; l=1}^{k}\{X_j > u\}\{X_l > u\} \,. \qquad (III.1)$$

Proof. The proof of the first inequality is trivial, namely if the left hand side is zero, the right hand side also. On the other hand the right hand side is less than or equal to one and the left hand side is larger than or equal to one if it is not zero.

For the second inequality, we have

$$\{Z_k > u\} \geq \sum_{j=1}^{k}\{X_j > u\} \prod_{l\neq j\; l=1}^{k}\{X_l < u\}$$

i.e. if only one $X_j > u$ then the sup is obviously larger than u. Therefore

$$\{Z_k > u\} \geq \sum_{j=1}^{k}\{X_j > u\} - \sum_{j=1}^{k}\{X_j > u\}\left(1 - \prod_{l\neq j\; l=1}^{k}\{X_l < u\}\right) \,.$$

On the other hand, as in the first inequality we have

$$1 - \prod_{l\neq j\; l=1}^{k}\{X_l < u\} \leq \sum_{l\neq j\; l=1}^{k}\{X_l > u\} \,,$$

and this implies the lower bound.

Proposition III.2. *For any integers s, r, m, k, $p \geq 0$ we have*

$$0 \leq \mathbb{P}(Z_r < u) - \mathbb{P}(Z_{r+k} < u) \leq k\mathbb{P}(X > u) \,.$$

and

$$\left|\mathbb{P}(Z_{m+p+s} < u) - \mathbb{P}(Z_m < u) + \sum_{j=1}^{p}\mathbb{E}(\{X > u\}\{Z_m < u\}\circ f^{p+s-j})\right| \leq$$

$$2p\sum_{j=1}^{p}\mathbb{P}(\{X > u\}\{X > u\}\circ f^{j}) + s\mathbb{P}(X > u) \,.$$

Proof. We have of course

$$0 \leq \mathbb{P}(Z_r < u) - \mathbb{P}(Z_{r+k} < u) \leq \sum_{j=0}^{k-1}\left(\mathbb{P}(Z_{r+j} < u) - \mathbb{P}(Z_{r+j+1} < u)\right) \,.$$

On the other hand, for any $l \geq 0$

$$\mathbb{P}(Z_l < u) = \mathbb{P}(Z_{l+1} < u) + \mathbb{P}(Z_l < u,\ X_{l+1} > u) \leq \mathbb{P}(Z_{l+1} < u) + \mathbb{P}(X_{l+1} > u)$$

and the first result follows by stationarity.
We now observe that

$$\{Z_{m+p+s} < u\} = \{Z_p < u\}\ \{Z_s < u\} \circ f^p\ \{Z_m < u\} \circ f^{p+s}\ .$$

It follows easily from this identity that

$$|\{Z_{m+p+s} < u\} - \{Z_p < u\}\ \{Z_m < u\} \circ f^{p+s}| \leq \{Z_s > u\} \circ f^p\ .$$

Therefore, using Lemma III.1 we get

$$\Big|\mathbb{E}(\{Z_{m+p+s} < u\}) - \mathbb{E}(\{Z_p < u\}\ \{Z_m < u\} \circ f^{p+s})\Big| \leq s\mathbb{P}(X > u)\ .$$

Using $\{Z_p < u\} = 1 - \{Z_p > u\}$, Lemma III.1 and stationarity, we get

$$\left|\mathbb{E}(\{Z_p < u\}\ \{Z_m < u\} \circ f^{p+s}) - \mathbb{E}(\{Z_m < u\}) + \sum_{j=1}^{p} \mathbb{E}(\{X > u\}\{Z_m < u\} \circ f^{p+s-j})\right| \leq$$

$$2p \sum_{j=1}^{p} \mathbb{P}(\{X > u\}\{X > u\} \circ f^j)\ ,$$

and the result follows.

The decay of correlations is always used below in the same form, and we present this estimate independently. It is formulated in terms of the rate of decay α_ω for Hölder continuous functions of exponent ω.

Lemma III.3. *For any positive number η, for any integer s and for any interval I and any set A, we have*

$$\Big|\mathbb{P}(I \cap f^{-s}(A)) - \mathbb{P}(I)\mathbb{P}(A)\Big| \leq |I|^{-\omega(1+\eta)}\alpha_\omega(s) + \mathcal{O}(1)|I|^{\theta(1+\eta)}\ ,$$

where θ is the number given in Lemma II.2.

Proof. The decay of correlations is formulated for Hölder continuous functions in [Y2], and does not apply as such to characteristic functions. However, if I is an interval, for any number $\eta > 0$ we can find a function ϕ which is non negative, satisfies $\phi \leq \chi_I$, is Lipschitz with a Lipshitz constant smaller than $|I|^{-1-\eta}$ and such that the support of $\chi_I(1-\phi)$ is within a distance $|I|^{1+\eta}$ of the boundary of I (take for example the linear interpolation).

We now apply the decay of correlations for functions which are Hölder continuous with exponent ω and get

$$\left|\int \phi\ \chi_A \circ\ f^s\ d\mu - \int \phi\ d\mu \int \chi_A\ d\mu\right| \leq |I|^{-\omega(1+\eta)}\alpha_\omega(s)\ .$$

Using now Lemma II.2 we get

$$\Big|\mathbb{P}(I\cap f^{-s}(A))-\mathbb{P}(I)\mathbb{P}(A)\Big|\le |I|^{-\omega(1+\eta)}\alpha_\omega(s)+\mathcal{O}(1)|I|^{\theta(1+\eta)}\ .$$

Remark. The decay of correlations in [Y2] is not really formulated for Hölder continuous functions but in term of estimates using the function s. It is easy to show that any Hölder continuous function u of Hölder exponent ω satisfies these estimates.

We now review and collect all the estimates. We start by defining the set of full measure for which Theorem I.1 holds. This is the set of x for which

$$\lim_{a\to 0}\frac{1}{2a}\mu([x-a,x+a])=h(x)$$

and which belong to only finitely many sets F_k defined in section 2. It follows from the Lebesgue differentiation theorem applied to $\mu=h\,d\lambda$ (see for example [St]) and Lemma II.5 that the above two properties hold for a set of full measure.

For a fixed $v>0$ define the sequence (u_n) by

$$u_n=v+\log n\ .$$

Let $k(x)$ be the smallest integer such that $x\notin F_j$ for any $j\ge k(x)$. From now on, we will assume $n>3(1+e^{-v})k(x)^{2\psi}$ where ψ is the constant appearing in Lemma II.5.

We define the integer p by $p=[\sqrt{n}]$ where $[\,\cdot\,]$ denotes the integer part. The integers q and r are given by the Euclidean division of n by p, $n=pq+r$ and $0\le r<p$. Finally we define $s=[\log n]^2$. These choices are only made for definiteness. These choices for the numbers p, q and r are only convenient ones. many other choices work as well.

We now replace $\mathbb{P}(Z_n<u_n)$ by $\mathbb{P}(Z_{q(p+s)}<u_n)$ and by Proposition III.2 this produces an error at most

$$|\mathbb{P}(Z_n<u_n)-\mathbb{P}(Z_{q(p+s)}<u_n)|\le qs\mathbb{P}(X>u_n)\ .$$

We now estimate recursively the numbers $\mathbb{P}(Z_{l(p+s)}<u_n)$ for $0\le l\le q$. Using Lemmata III.2 and III.3 we have for any $q\ge l\ge 1$

$$|\mathbb{P}(Z_{l(p+s)}<u_n)-(1-p\mathbb{P}(X>u_n))\mathbb{P}(Z_{(l-1)(p+s)}<u_n)|\le\Gamma_n$$

where

$$\Gamma_n=s\mathbb{P}(X>u_n)+2p\sum_{j=1}^{p}\mathbb{P}\left(\big\{X>u_n\big\}\cap\big\{X>u_n\big\}\circ f^j\right)$$

$$+p|\{X>u_n\}|^{-\omega(1+\eta)}\alpha_\omega(s)+p\,\mathcal{O}(1)\,|\{X>u_n\}|^{\theta(1+\eta)}\ .$$

We finally get if $p\mathbb{P}(X>u_n)<2$

$$\Big|\mathbb{P}(Z_{q(p+s)}<u_n)-(1-p\mathbb{P}(X>u_n))^q\Big|\le q\Gamma_n\ .$$

From Lebesgues differentiation theorem we have

$$\lim_{n\to\infty} pq\mathbb{P}(X > u_n) = 2e^{-v}h(x)$$

and since s/p tends to zero when n tends to infinity,

$$\lim_{n\to\infty} qs\mathbb{P}(X > u_n) = 0 \ .$$

A similar argument ensures $p\mathbb{P}(X > u_n) < 2$ for n large enough. In order to finish the proof of Theorem I.1, we have to show that

$$\lim_{n\to\infty} q\Gamma_n = 0 \ .$$

If we chose η such that $\theta(1+\eta) > 2$, the result is obvious using the exponential decay of α_ω except for the term

$$qp\sum_{j=1}^{p} \mathbb{P}\left(\{X > u_n\} \cap \{X > u_n\} \circ f^j\right) \ .$$

Using the decay of correlations, we have easily

$$qp\sum_{j=s}^{p} \mathbb{P}\left(\{X > u_n\} \cap \{X > u_n\} \circ f^j\right)$$

$$\leq qp^2\mathbb{P}(X > u_n)^2 + qp^2|\{X > u_n\}|^{-\omega(1+\eta)}\alpha_\omega(s) + qp^2\,\mathcal{O}(1)\,|\{X > u_n\}|^{\theta(1+\eta)} \ .$$

With the above choice of η and the exponential decay of α_ω, this term tends to zero when n tends to infinity. It remains to control the part of the above sum running from $j = 1$ to $j = s - 1$.

We now define an integer k (which depends on n) by

$$k = \left[\left(ne^v/3\right)^{1/\psi}\right] \ .$$

Recall that n is large enough so that x does not belong to F_k. We now observe from the definitions that for $j \leq s$ (and for n large enough)

$$\{X > u_n\} \cap \{X > u_n\} \circ f^j \subset [x - k^{-\psi}, x + k^{-\psi}] \cap E_{k^\psi} \ .$$

Since $x \notin F_k$ this implies

$$\mathbb{P}\left(\{X > u_n\} \cap \{X > u_n\} \circ f^j\right) \leq \mathcal{O}(1)k^{-\psi(1+\rho)} \leq \mathcal{O}(1)n^{-(1+\rho)} \ .$$

We finally get a bound

$$qp\sum_{j=1}^{s} \mathbb{P}\left(\{X > u_n\} \cap \{X > u_n\} \circ f^j\right) \leq \mathcal{O}(1)\frac{qps}{n^{1+\rho}} \ ,$$

which tends to zero when n tends to infinity. This finishes the proof of Theorem I.1.

IV. STATISTICS OF NEAREST RECURRENCE.

In this section we discuss a variant of Theorem I.1 which gives the fluctuations for the nearest return to the starting point. We define a sequence of real valued random variables (X_j) by

$$X_j(x) = -\log d(x, f^j(x)) \ .$$

We then define the sequence of random variables (Z_n) by

$$Z_n(x) = \sup_{1\le j\le n} X_j(x) \ ,$$

and ask if the sequence of random variables $(Z_n - \log n)$ converges in law. This is indeed the case under the same hypothesis as in Theorem I.1.

Theorem IV.1. *For maps of the interval satisfying the hypothesis* **H1**-**H4**, *and such that $\lambda(R > k)$ decays exponentially fast we have*

$$\lim_{n\to\infty} \mathbb{P}\left(Z_n < s + \log n\right) = \int e^{-2e^{-s}h(x)} h(x)\, dx$$

where h is the density of the invariant measure.

Note that here also the normalization is related to the dimension of the measure. In more general cases one may also expect to obtain log-normal fluctuations as in [C.G.S.] and [K.] instead of an exponential law.

The proof is similar to that of Theorem I.1 except that we have to use the decay of correlations to separate the initial constraint. We will explain in details how this can be done, and leave to the reader to reproduce the part of the argument which is identical to the proof of Theorem I.1.

Proof. For a given integer n, let $\mathcal{U}_n$ be the uniform partition of the interval $[a, b]$ by intervals of length $1/n^{1+\beta'/10}$ where β' is the exponent appearing in Corollary II.4 (the last segment being of length at most this number). We fix a positive number v, and from now on we will assume that $n > (1 + e^v)^2$. If $\Delta \in \mathcal{U}_n$, we define two intervals Δ^+ and Δ^- by

$$\Delta^\pm = \left\{ x \,|\, d(x, \Delta) \le n^{-1}e^{-v} \pm n^{-1-\beta'/10} \right\} \ .$$

With this notation, we have obviously

$$\sum_{\Delta\in\mathcal{U}_n} \mathbb{P}\left(\Delta\, ,\, f^j(\,\cdot\,) \notin \Delta^-\, ,\, j = 1, \cdots, n\right)$$

$$\ge \mathbb{P}\left(Z_n < v + \log n\right) \ge$$

$$\sum_{\Delta\in\mathcal{U}_n} \mathbb{P}\left(\Delta\, ,\, f^j(\,\cdot\,) \notin \Delta^+\, ,\, j = 1, \cdots, n\right) \ .$$

We define $p = [n^{\theta/2}]$ (θ as given in Lemma II.2), $s = [(\log n)^2]$ and let $n = (p+s)q+r$ with $0 \le r < p + s$ be the division of n by $p + s$. As in the first step of the proof of Theorem I.1, we wish to replace n by $q(p + s)$.

We have obviously

$$\left|\mathbb{P}\left(\Delta\,,\, f^j(\,\cdot\,)\notin\Delta^+\,,\, j=1,\,\cdots\,,\, n\right)-\mathbb{P}\left(\Delta\,,\, f^j(\,\cdot\,)\notin\Delta^+\,,\, j=1,\,\cdots\,,\,(p+s)q\right)\right|\le$$

$$\sum_{j=q(p+s)+1}^{j=q(p+s)+r}\mathbb{P}\left(\Delta\,,\, f^j(\,\cdot\,)\in\Delta^+\right)\ .$$

Using decay of correlations as in Lemma III.3, we choose $\eta>3/\theta$ and the above quantity is bounded by

$$\mathcal{O}(1)\, r\,\left(\mu(\Delta)\mu(\Delta^+)+n^{2\omega(1+\eta)}\alpha_\omega(pq)+n^{-3-\theta}\right)\ .$$

Using Lemma II.2, the first term is bounded by

$$\mathcal{O}(1)\, r\,\mu(\Delta)n^{-\theta}\ ,$$

and since $r\le p+s\le 2n^{\theta/2}$, we can sum this quantity over Δ and get a bound $\mathcal{O}(1)n^{-\theta/3}$. For the two other terms, we use the fact that the cardinality of $\mathcal{U}_n$ is $\mathcal{O}(1)n^2$ and $\alpha_\omega(pq)$ decays exponentially fast in n. Note that this above bounds may not apply to the last Δ in $\mathcal{U}_n$ which may be of size much smaller than $n^{-1-\beta'/10}$. The reader can easily convince himself that this segment will contribute at most $\mathcal{O}(1)n^{-\theta(1+\beta'/10)}$ to the final result. A similar estimate holds for the terms involving Δ^- instead of Δ^+.

For $\Delta\in\mathcal{U}_n$ we define a set $B^+_{p,q,s}(\Delta)$ by

$$B^+_{p,q,s}(\Delta)=\left\{x\in\Delta\,|\, f^j(x)\notin\Delta^+\ 1\le j\le(p+s)q\right\}\ ,$$

and similarly for $B^-_{p,q,s}(\Delta)$. From the previous bound, we now have to estimate

$$\sum_{\Delta\in\mathcal{U}_n}\mathbb{P}(B^\pm_{p,q,s}(\Delta))\ .$$

We will now eliminate the constraint $x\in\Delta$ in the definition of $B^\pm_{p,q,s}(\Delta)$. For a fixed $v\in\mathbf{R}$, we assume from now on n large enough so that

$$e^{-(\log n)^{1/2}}>\frac{2e^{-v}}{n}+\frac{2}{n^{1+\beta'/10}}\ .$$

Let

$$G_n=\left\{x\;\Big|\;\forall\, 1\le j\le(\log n)^2\,,\ |x-f^j(x)|\ge e^{-\sqrt{\log n}}\right\}\ ,$$

this definition implies that if $x\in\Delta\cap G_n$, we have

$$f^j(x)\notin\Delta^+\quad\text{for}\quad j=1,\,\cdots\,,\,\left[(\log n)^2\right]\ .$$

Therefore, if we define $\tilde{B}^+_{p,q,s}(\Delta)\supset B^+_{p,q,s}(\Delta)$ by

$$\tilde{B}^+_{p,q,s}(\Delta)=\left\{x\in\Delta\,|\, f^j(x)\notin\Delta^+\ s\le j\le(p+s)q\right\}\ ,$$

we have

$$G_n \cap B^+_{p,q,s}(\Delta) = G_n \cap \tilde{B}^+_{p,q,s}(\Delta) \ .$$

Therefore

$$\Big|\mathbb{P}(B^+_{p,q,s}(\Delta)) - \mathbb{P}(\tilde{B}^+_{p,q,s}(\Delta))\Big| \le \mathbb{P}\big(\Delta \cap G^c_n\big) \ .$$

The sum over Δ of this quantity is equal to $\mu(G^c_n)$. However

$$G^c_n \subset E_{e^{(\log n)^{1/2}}}$$

which implies that $\mu(G^c_n)$ tends to zero when n tends to infinity by Corollary II.4. It is therefore enough to estimate $\mathbb{P}(\tilde{B}^+_{p,q,s}(\Delta))$.

We now use the decay of correlations from Lemma III.3 and the estimate $\mu(\Delta^+) \le \mathcal{O}(1)n^{-\theta}$ from Lemma II.2 to replace $\mu(\tilde{B}^+_{p,q,s}(\Delta))$ by $\mu(\Delta)\mu(C^+_{p,q,s}(\Delta))$ where $C^+_{p,q,s}(\Delta)$ is defined by

$$C^+_{p,q,s}(\Delta) = \big\{x \,|\, f^j(x) \notin \Delta^+ \ , \ 0 \le j \le (p+s)q\big\} \ .$$

The proof then proceeds following the proof of Theorem I.1 provided $\mu(\Delta^+)$ is small enough, for example we can take $\mu(\Delta^+) \le |\Delta^+| \log n$. This is needed in order to estimate as in section III the first part of the remainder term

$$\mathcal{O}(1) \sum_{\Delta} \in \mathcal{U}_n \mu(\Delta) \left[qs\mu(\Delta^+) + qp^2\mu(\Delta^+)^2\right] \ .$$

For the second part of the remainder term, instead of using the sequence of sets (F_k) as in section III, one can define for each integer n a subset $\mathcal{F}^+_n$ of $\mathcal{U}_n$ by

$$\mathcal{F}^+_n = \Big\{\Delta \in \mathcal{U}_n \mid \mu\big(\Delta^+ \cap E_{[n^{2/3}]}\big) \ge n^{-1-\beta'/2}\Big\}$$

where β' is the constant appearing in Corollary II.4. We have by Corollary II.4

$$C' n^{-2\beta'/3} \ge \mu\big(E_{[n^{2/3}]}\big) \ge \frac{n^{-\beta'/10}}{4(1+e^{-v})} \sum_{\Delta \in \mathcal{F}^+_n} \mu\big(\Delta^+ \cap E_{[n^{2/3}]}\big) \ge \frac{n^{-\beta'/10}}{4(1+e^{-v})} n^{-1-\beta'/2} \# \mathcal{F}^+_n \ ,$$

where $\#$ denotes the cardinality. The factor $n^{-\beta'/10}/4$ comes from the fact that the sets Δ^+ are not disjoint, but if we take every other $n^{\beta'/10}e^{-v}$ such sets, we get a disjoint collection for n large enough. Therefore

$$\mu\left(\bigcup_{\Delta \in \mathcal{F}^+_n \ , \ \mu(\Delta^+) \le |\Delta^+| \log n} \Delta\right) \le$$

$$\mu\left(\bigcup_{\Delta \in \mathcal{F}^+_n \ , \ \mu(\Delta^+) \le |\Delta^+| \log n} \Delta^+\right) \le \mathcal{O}(1) n^{-1} \# \mathcal{F}^+_n \log n \le \mathcal{O}(1) n^{-\beta'/20}$$

which tends to zero when n tends to infinity. We finally get

$$\mathbb{P}\left(Z_n < s + \log n\right) \geq \sum_{\Delta \in \mathcal{U}_n\,,\,\mu(\Delta^+) \leq |\Delta^+| \log n} \mu(\Delta)\; e^{-n\mu(\Delta^+)} - o(1)\;.$$

A similar upper bound follows with Δ^- instead of Δ^+, although with an additional term, namely

$$\mathbb{P}\left(Z_n < s + \log n\right) \leq \sum_{\Delta \in \mathcal{U}_n\,,\,\mu(\Delta^-) \leq |\Delta^-| \log n} \mu(\Delta)\; e^{-n\mu(\Delta^-)} + o(1) + \sum_{\Delta \in \mathcal{U}_n\,,\,\mu(\Delta^-) > |\Delta^-| \log n} \mu(\Delta)\;.$$

By Lebesgue's derivation theorem and dominated convergence theorem, we deduce that

$$\lim_{n\to\infty} \sum_{\Delta \in \mathcal{U}_n} \mu(\Delta)\; e^{-n\mu(\Delta^\pm)} = \int e^{-2e^{-v}h(x)} h(x)dx\;.$$

It remains to control the sum of the measure of the elements Δ of $\mathcal{U}_n$ such that $\mu(\Delta^\pm) > |\Delta^\pm| \log n$.

By Lemma II.2, it follows that the density h of μ belongs to some L^σ with $\sigma > 1$. Therefore, from the maximal theorem of Hardy and Littlewood [St.] it follows that the maximal function

$$Mh(x) = \sup_{a>0} \frac{1}{2a} \int_{x-a}^{x+a} h(y)dy$$

also belongs to L^σ. For $\rho > 0$, let D_ρ be the set

$$D_\rho = \{x \,|\, Mh(x) > \rho\} = \left\{x \,|\, \sup_{a>0} a^{-1} \mu([x-a, x+a] > 2\rho\right\}\;.$$

We have by Chebychev's inequality

$$\lambda(D_\rho) \leq \mathcal{O}(1)\; \rho^{-\sigma}\;.$$

We now observe that if for a $\Delta \in \mathcal{U}_n$ we have $\mu(\Delta^+) \geq |\Delta^+| \log n$, then for any $y \in \Delta$ we have (for n large enough)

$$\begin{aligned} \mu([y-(1+e^{-v})n^{-1}, y+(1+e^{-v})n^{-1}]) &\geq\; \mu(\Delta^+) \;\geq\; 2e^{-v}n^{-1}\log n \\ &\geq \left|[y-(1+e^{-v})n^{-1}, y+(1+e^{-v})n^{-1}]\right| (\log n)^{1/2}\;, \end{aligned}$$

namely $y \in D_{(\log n)^{1/2}}$ for any $y \in \Delta$, hence $\Delta \subset D_{(\log n)^{1/2}}$. Therefore

$$\lambda\left(\bigcup_{\Delta\,,\,\mu(\Delta^+) > \log n\, |\Delta^+|} \Delta\right) = \sum_{\Delta\,,\,\mu(\Delta^+) > \log n\, |\Delta^+|} \lambda(\Delta) \leq \lambda\big(D_{(\log n)^{1/2}}\big)\;,$$

which tends to zero when n tends to infinity. We now use Lemma II.2 to conclude that

$$\lim_{n\to\infty} \mu\left(\bigcup_{\Delta\,,\,\mu(\Delta^+) > \log n\, |\Delta^+|} \Delta\right) = 0\;.$$

A similar argument holds for the case of Δ^-, one can also observe that $\mu(\Delta^+) \leq |\Delta^+| \log n$ implies for n large enough $\mu(\Delta^-) \leq 2|\Delta^-| \log n$. This completes the proof of Theorem IV.1.

References.

[ABST] H.Abarbanel, R.Brown, J.Sidorowich, and L.Tsimring, The Analysis of Observed Chaotic Data in Physical Systems. Reviews of Mod. Phys. **65**, 1331-1392 (1993).

[BV] V.Baladi, M.Viana. Strong stochastic stability and rate of mixing for unimodal maps. Ann. Sci. Ec. Norm. Super. **29**, 483-517 (1996).

[CGS] P.Collet, A.Galves and B.Schmitt. Fluctuations of Repetition Times for Gibbsian Sources. Preprint.

[G] J.Galambos. *The asymptotic theory of extreme order statistics*. Wiley Series in Probability and Mathematical Statistics, New York (1978).

[GS] A.Galves, B.Schmitt. Inequalities for hitting times in mixing dynamical systems. Random Comput. Dyn. **5**, 337-348 (1997).

[K] I.Kontoyiannis. Asymptotic Recurrence and waiting times for stationary processes. J. Theoret. Probab. to appear.

[KN] G.Keller, T.Nowicki. Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps. Commun. Math. Phys. **149**, 31-69 (1992).

[Sh] P. C. Shields. The Interactions Between Ergodic Theory and Information Theory IEEE Trans. on Information Theory. **44**, 2079-2093 (1998).

[St] E.Stein. *Harmonic analysis: Real-variable methods, orthogonality, and oscillatory integrals*. Princeton Mathematical Series. 43. Princeton, NJ: Princeton University Press, 1993.

[T] M.Thaler. Transformations on $[0, 1]$ with infinite invariant measures. Isr. J. Math. **46**, 67-96 (1983).

[WZW] A. D. Wyner, J. Ziv, and A. J. Wyner. On the Role of Pattern Matching in Information Theory. IEEE Trans. on Information Theory. **44**, 2045-2056 (1998).

[Y1] L.S.Young. Statistical properties of dynamical systems with some hyperbolicity. Ann. Math. **147**, 585-650 (1998).

[Y2] L.S.Young. Recurrence times and rates of mixing. Isr. J. Math. to appear.